ELSEVIER Linear Algebra and its Applications 359 (2003) 263–276

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

A $\sin 2\Theta$ theorem for graded indefinite Hermitian matrices ☆

Ninoslav Truhar [a,*], Ren-Cang Li [b]

[a] *Lehrgebiet Mathematische Physik, Fernuniversität, 58084 Hagen, Germany*
[b] *Department of Mathematics, University of Kentucky, Lexington, KY 40506, USA*

Received 20 January 2001; accepted 10 May 2002

Submitted by B.N. Parlett

Abstract

This paper gives double angle theorems that bound the change in an invariant subspace of an indefinite Hermitian matrix in the graded form $H = D^*AD$ subject to a perturbation $H \to \widetilde{H} = D^*(A + \Delta A)D$. These theorems extend recent results on a definite Hermitian matrix in the graded form (Linear Algebra Appl. 311 (2000) 45) but the bounds here are more complicated in that they depend on not only relative gaps and norms of ΔA as in the definite case but also norms of some J-unitary matrices, where J is diagonal with ± 1 on its diagonal. For two special but interesting cases, bounds on these J-unitary matrices are obtained to show that their norms are of moderate magnitude.

Keywords: Relative perturbation bounds; Invariant subspaces

1. Introduction

Let H and $\widetilde{H}$ be two Hermitian matrices whose eigen-decompositions are

$$H = [U_1 \quad U_2] \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix},$$

☆ This report is available on the web at http://www.ms.uky.edu/~rcli/. This work was supported in part by the National Science Foundation under grant no. ACI-9721388 and by the National Science Foundation CAREER award under grant no. CCR-9875201.

* Corresponding author. Address: Faculty of Civil Engineering, University of Osijek, Drinska 16a, HR-3100, Osijek, Croatia.

E-mail addresses: truhar@gfos.hr, ninoslav.truhar@fernuni-hagen.de (N. Truhar), rcli@ms.uky.edu (R.-C. Li).

0024-3795/02/$ - see front matter
PII: S0024-3795(02)00424-X

$$\widetilde{H} = [\widetilde{U}_1 \quad \widetilde{U}_2] \begin{bmatrix} \widetilde{\Lambda}_1 & 0 \\ 0 & \widetilde{\Lambda}_2 \end{bmatrix} \begin{bmatrix} \widetilde{U}_1^* \\ \widetilde{U}_2^* \end{bmatrix}, \tag{1.1}$$

where $U = [U_1 \quad U_2]$, $\widetilde{U} = [\widetilde{U}_1 \quad \widetilde{U}_2]$ are unitary, and

$$\Lambda_1 = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k), \quad \Lambda_2 = \mathrm{diag}(\lambda_{k+1}, \lambda_{k+2}, \ldots, \lambda_n), \tag{1.2}$$

$$\widetilde{\Lambda}_1 = \mathrm{diag}(\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_k), \quad \widetilde{\Lambda}_2 = \mathrm{diag}(\widetilde{\lambda}_{k+1}, \widetilde{\lambda}_{k+2}, \ldots, \widetilde{\lambda}_n). \tag{1.3}$$

We are interested in bounding the changes in subspace $\mathscr{S} \stackrel{\text{def}}{=} \mathrm{span}(U_1)$, H's invariant subspace spanned by U_1's columns. We shall do this by bounding the sines of the double canonical angles between $\mathscr{S}$ and $\widetilde{\mathscr{S}} \stackrel{\text{def}}{=} \mathrm{span}(\widetilde{U}_1)$. For *absolute* perturbations, i.e., $H \to H + \Delta H \equiv \widetilde{H}$, this was done by Davis and Kahan [1], and for *multiplicative* perturbations, i.e., $H \to D_M^* H D_M \equiv \widetilde{H}$, as well as for perturbations involving graded *definite* Hermitian matrices, this was done by [4], where D_M is assumed close to the identity matrix. We say Hermitian matrix $H = D^*AD$ is in the *graded form* if A is well-conditioned, i.e., $\|A\|\|A^{-1}\|$ is of moderate magnitude while $\|H\|\|H^{-1}\| \gg \|A\|\|A^{-1}\|$. Usually D, the scaling matrix, is diagonal; but our theorems in general do not require this unless it says so. A perturbed graded Hermitian matrix takes the form $\widetilde{H} = D^*(A + \Delta A)D$.

In the case of absolute perturbations, (absolute) perturbation bounds on the changes in $\mathscr{S}$ are typically proportional to the norm of ΔH and to the reciprocal of the difference (called the *absolute gap*) $\min |\lambda_i - \widetilde{\lambda}_j|$ over $i \leqslant k$ and $j > k$. Such bounds may be useless for subspaces corresponding with eigenvalues of tiny magnitude that are perfectly distinguishable from the rest but associated absolute gaps are tiny. This problem, nonetheless, disappears if gaps were measured as if all eigenvalues had the same exponent 0. That is exactly what relative perturbation theories (see, e.g., [4,8] and reference therein) developed for multiplicative perturbations and perturbations in the graded cases attempt to do and accomplish. The main results of this paper are extensions of a $\sin 2\Theta$ theorem for *positive definite* Hermitian matrices [4] to *indefinite graded* Hermitian matrices. An advantage of a double angle theorem over a single angle theorem is that a double angle theorem uses a relative gap involving the spectrum of only one matrix, either H or $\widetilde{H}$, in contrast to the gap used by single angle theorems.

Notation. $\|X\|$ and $\|X\|_{\mathrm{F}}$ are the spectral and Frobenius norms of matrix X, respectively. X^* is the conjugate transpose, and $\lambda(X)$ is the spectrum of X. I_n denotes the $n \times n$ identity matrix (we may simply write I instead if no confusion).

2. Hyperbolic singular value decomposition

Hyperbolic singular value decomposition (HSVD) [5,9] provides an important tool in our later developments. Throughout this section, Z is $n \times n$ and nonsingular, and J is $n \times n$ and diagonal with ± 1 on its main diagonal.

Theorem 2.1 (*HSVD*). *There exist an $n \times n$ unitary matrix Y and an $n \times n$ nonsingular matrix X such that*

$$Z = Y\Sigma X^{-1}, \quad X^*JX = J, \tag{2.1}$$

where Σ is $n \times n$ and diagonal with positive real diagonal entries.

We call (2.1) the *HSVD of Z* (*with respect to J*). HSVD theorem in its generality allows Z to be rectangular, but the square case is what we will need in this paper.

A matrix X is called *J-unitary* if $X^*JX = J$. It can be proved that *if X is J-unitary, so are X^*, X^{-1}, and X^{-*}*. In fact $X^*JX = J$ implies immediately $X^{-*}JX^{-1} = J$ and thus X^{-1} is J-unitary. Note also $X = JX^{-*}J$ to get

$$(X^*)^*JX^* = XJX^* = JX^{-*}JJX^* = JX^{-*}X^* = J,$$

so X^* is J-unitary. Finally $X^{-*} = (X^*)^{-1}$ is J-unitary.

Let M be positive definite and $J = \mathrm{diag}(\pm 1)$. Matrix pair $\{M, J\}$ is *definite* since $|x^*Mx|^2 + |x^*Jx|^2 \geqslant |x^*Mx|^2 > 0$ for all vectors $x \neq 0$ [7, p. 318]. Thus there is a nonsingular matrix X whose columns consist of a complete set of (generalized) eigenvectors of the pair such that

$$X^*MX = \text{diagonal}, \quad X^*JX = J. \tag{2.2}$$

We call this X the *canonical eigenvector matrix* of the pair $\{M, J\}$. See [7, p. 318] for more about eigen-decompositions of definite matrix pairs.

HSVD of Z (with respect to J) is closely related to the canonical eigenvector matrix of the definite pair $\{Z^*Z, J\}$ and the eigen-decomposition of Hermitian ZJZ^*. This is given in the following theorem.

Theorem 2.2. *In Theorem* 2.1,

1. *X is the canonical eigenvector matrix of the pair $\{Z^*Z, J\}$.*
2. *$ZJZ^* = Y(\Sigma J\Sigma)Y^*$ is an eigen-decomposition.*

Proof. In Theorem 2.1, by (2.1) we have $X^*JX = J$ and also

$$X^*Z^*ZX = \underbrace{X^*X^{-*}}\Sigma\underbrace{Y^*Y}\Sigma\underbrace{X^{-1}X} = \Sigma\Sigma = \Sigma^2,$$
$$ZJZ^* = Y\Sigma X^{-1}JX^{-*}\Sigma Y^* = Y\Sigma J\Sigma Y^*.$$

These complete the proof. □

On the other hand, ZJZ^* is Hermitian. Thus there is a unitary matrix Y such that $ZJZ^* = Y\Lambda Y^*$, where Λ is diagonal. Because Z is nonsingular, J and Λ have the same inertia, i.e., the same number of positive diagonal entries and the same number of negative diagonal entries. Therefore the diagonal entries of Λ and accordingly the columns of Y can be ordered so that $\Lambda = |\Lambda|J$. Define $\Sigma = |\Lambda|^{1/2}$ and $X = Z^{-1}Y\Sigma$.

It can be verified directly that $X^*JX = J$ and $Z = Y\Sigma X^{-1}$, the HSVD of Z. In fact, $Z = Y\Sigma X^{-1}$ is evident from the definition of X, and that X is J-unitary is proved as follows:

$$\begin{aligned} X^*JX &= \Sigma Y^* Z^{-*} J Z^{-1} Y\Sigma = \Sigma Y^*(ZJZ^*)^{-1} Y\Sigma \\ &= \Sigma Y^* Y \Lambda^{-1} Y^* Y \Sigma = \Sigma\Lambda^{-1}\Sigma = J. \end{aligned}$$

3. Main results

Let H and $\widetilde{H}$ and their eigen-decompositions be as described in Section 1. Similarly, as in [4] we shall define a unitary matrix

$$S = [U_1 \quad U_2]\begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{bmatrix}\begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix} = U_1U_1^* - U_2U_2^*. \tag{3.1}$$

Note that

$$S^* = S, \quad S^2 = I_n, \quad S^{-1} = S, \quad SHS = H.$$

We now define an auxiliary matrix $\widehat{H}$ as

$$\widehat{H} \equiv S\widetilde{H}S = [\widehat{U}_1 \quad \widehat{U}_2]\begin{bmatrix} \widetilde{\Lambda}_1 & \\ & \widetilde{\Lambda}_2 \end{bmatrix}\begin{bmatrix} \widehat{U}_1^* \\ \widehat{U}_2^* \end{bmatrix}, \tag{3.2}$$

where $\widehat{U}_i = S\widetilde{U}_i$ for $i = 1, 2$. Geometrically, $S\widetilde{\mathscr{S}}$ is a reflection of $\widetilde{\mathscr{S}}$ where the mirror for S is $\mathscr{S}$ and S reverses $\mathscr{S}^\perp$, the orthogonal complement of $\mathscr{S}$, as shown in Fig. 1.

This explains the following equation due to Davis and Kahan [1] (see also [4]):

$$\left\|\sin\Theta\big(\widetilde{U}_1, \widehat{U}_1\big)\right\|_{\mathrm{F}} = \left\|\sin 2\Theta\big(U_1, \widetilde{U}_1\big)\right\|_{\mathrm{F}}. \tag{3.3}$$

In what follows we shall seek bounds for $\left\|\sin\Theta\big(\widetilde{U}_1, \widehat{U}_1\big)\right\|_{\mathrm{F}}$.

Since the positive semidefinite case has been studied by [4], interesting to us is the case when H is indefinite Hermitian in the graded form $H = D^*AD$. Suppose H is perturbed to

$$\widetilde{H} = D^*\widetilde{A}D \equiv D^*(A + \Delta A)D.$$

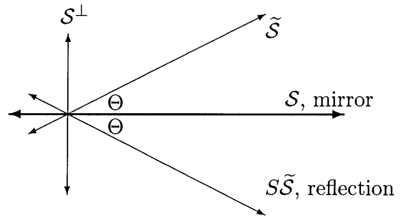

Fig. 1.

Let A's eigen-decomposition be

$$A = Q\Omega Q^* = Q|\Omega|^{1/2}J|\Omega|^{1/2}Q^*, \tag{3.4}$$

where J is diagonal with ± 1 on its main diagonal. It can be seen that the diagonal elements of J are the signs of the corresponding eigenvalues of A. Sylvester's theorem [2, Theorem 4.5.8] implies A, H, and J all have the same inertia. Set

$$G = D^* Q|\Omega|^{1/2}, \tag{3.5}$$

$$E = |\Omega|^{-1/2}Q^*\Delta A Q|\Omega|^{-1/2} \tag{3.6}$$

to get

$$H = GJG^*, \quad \widetilde{H} = G(J+E)G^*. \tag{3.7}$$

We shall assume that $\|A^{-1}\|\|\Delta A\| < 1$ which insures $\|E\| \leqslant \|A^{-1}\|\|\Delta A\| < 1$ and the existence of $(I+EJ)^{1/2}$ defined by the following series [3, Theorem 6.2.8]

$$\begin{aligned} T &\stackrel{\text{def}}{=} (I+EJ)^{1/2} \\ &= I + \sum_{n=1}^{\infty}(-1)^{n-1}\frac{(2n-1)!!}{2^n n!}(EJ)^n, \end{aligned} \tag{3.8}$$

where $(2n-1)!! = 1\cdot 3\cdot 5\cdots(2n-1)$. It can be verified that $T = JT^*J$, and therefore

$$J + E = TJT^*, \tag{3.9}$$

$\|E\| < 1$ implies that $H = GJG^*$, $J+E$ and $\widetilde{H} = G(J+E)G^*$ all have the same inertia as J. From now on we will *assume that Λ_1 and $\widetilde{\Lambda}_1$ have the same inertia*, and thus (1.1) can be rewritten as

$$H = U|\Lambda|JU^*, \quad \widetilde{H} = \widetilde{U}|\widetilde{\Lambda}|J\widetilde{U}^*, \tag{3.10}$$

(if necessary some reordering may be needed for the columns of U_i and of $\widetilde{U}_i$, and the diagonal entries of Λ_i and $\widetilde{\Lambda}_i$ *without* affecting the invariant subspaces span(U_i) and span$(\widetilde{U}_i)$ in question), where $J = \operatorname{diag}(J_k, J_{n-k})$ is diagonal and partitioned conformably to (1.1). We have

$$H = GJG^*, \quad \widetilde{H} = GTJT^*G^* \equiv \widetilde{G}J\widetilde{G}^*, \quad \widetilde{G} = GT. \tag{3.11}$$

Bearing (3.10) and (3.11) in mind, we now invoke Theorem 2.2 to get the HSVDs of G and $\widetilde{G}$:

$$G = U|\Lambda|^{1/2}V^{-1}, \quad V^*JV = J, \tag{3.12}$$

$$\widetilde{G} = \widetilde{U}|\widetilde{\Lambda}|^{1/2}\widetilde{V}^{-1}, \quad \widetilde{V}^*J\widetilde{V} = J. \tag{3.13}$$

It is important to notice that the J's in (3.12) and (3.13) as well as the one in (3.17) below are the same because H, $\widetilde{H}$, and $\widehat{H}$ all have the same inertia and Λ_1 and $\widetilde{\Lambda}_1$

have the same inertia as guaranteed by the assumptions. To derive a bound for $\sin\Theta$ between $\widehat{U}_1$ and $\widetilde{U}_1$, we define

$$W \stackrel{\text{def}}{=} G^{-1}SG = V\begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k}\end{bmatrix}V^{-1}, \tag{3.14}$$

which is J-unitary, indeed

$$WJW^* = G^{-1}SGJG^*SG^{-*} = G^{-1}HG^{-*} = J,$$

where we have used $SHS = H$. Other properties of W useful to our later developments are

$$W^2 = I, \quad W^*J = JW, \quad \|W\| \leqslant \|V\|^2. \tag{3.15}$$

We use $GW = SG$ by (3.14) to get

$$\begin{aligned}\widehat{H} &= SGTJT^*G^*S = GT\,\underbrace{T^{-1}WT}\,J\,\underbrace{T^*W^*T^{-*}}\,T^*G^* \\ &= \widetilde{G}\widetilde{T}J\widetilde{T}^*\widetilde{G}^* \equiv \widehat{G}J\widehat{G}^*,\end{aligned}$$

where $\widetilde{G}$ is as in (3.11), and

$$\widetilde{T} = T^{-1}WT, \quad \widehat{G} = \widetilde{G}\widetilde{T}. \tag{3.16}$$

Similarly to (3.12) and (3.13) we can write the HSVD of $\widehat{G}$ as

$$\widehat{G} = \widehat{U}|\widetilde{\Lambda}|^{1/2}\widehat{V}^{-1}, \quad \widehat{V}^*J\widehat{V} = J. \tag{3.17}$$

It can be verified that

$$\widehat{H} - \widetilde{H} = \widehat{G}\big(J\widetilde{T^*} - \widetilde{T}^{-1}J\big)\widetilde{G}^*.$$

Pre- and post-multiply the equation by $\widehat{U}_2$ and $\widetilde{U}_1$, respectively, and use the fact that $\widehat{V}^{-1} = J\widehat{V}^*J$ and $\widetilde{V}^{-1} = J\widetilde{V}^*J$, to get

$$\widetilde{\Lambda}_2\widehat{U}_2^*\widetilde{U}_1 - \widehat{U}_2^*\widetilde{U}_1\widetilde{\Lambda}_1 = |\widetilde{\Lambda}_2|^{1/2}J_{n-k}\widehat{V}_2^*J\big(J\widetilde{T}^* - \widetilde{T}^{-1}J\big)J\widetilde{V}_1J_k|\widetilde{\Lambda}_1|^{1/2},$$

where $\widehat{V} = [\widehat{V}_1 \quad \widehat{V}_2]$ is partitioned accordingly. Now, we can state our main theorem, which generalizes [4, Theorem 2.3] to indefinite Hermitian matrices.

Theorem 3.1. *Let $H = D^*AD$ and $\widetilde{H} = D^*\widetilde{A}D$ be two $n \times n$ Hermitian matrices with eigen-decompositions* (1.1), *where A is nonsingular and satisfies $\|A^{-1}\|\|\Delta A\| < 1$, and let $J = \operatorname{diag}(\pm 1)$ be as determined in* (3.4). *Suppose Λ_1 and $\widetilde{\Lambda}_1$ have the same inertia. If*

$$\widetilde{\eta}_\chi \stackrel{\text{def}}{=} \min_{\mu\in\lambda(\widetilde{\Lambda}_1),\ \nu\in\lambda(\widetilde{\Lambda}_2)} \frac{|\mu-\nu|}{\sqrt{|\mu\nu|}} > 0,$$

then

$$\big\|\sin 2\Theta(U_1, \widetilde{U}_1)\big\|_{\mathrm{F}} \leqslant \|\widehat{V}_2\|\|\widetilde{V}_1\|\frac{\big\|J\widetilde{T}^* - \widetilde{T}^{-1}J\big\|_{\mathrm{F}}}{\widetilde{\eta}_\chi}, \tag{3.18}$$

where T and $\widetilde{T}$ are as in (3.8) *and* (3.16), $\widetilde{V}_1$ *and* $\widehat{V}_2$ *are from the HSVDs of* $\widetilde{G}$ *and* $\widehat{G}$ *in* (3.13) *and* (3.17), *associated with* $\widetilde{H}$ *and* $\widehat{H}$.

If A is positive definite, $J = I$ and $\|\widehat{V}_2\| = \|\widetilde{V}\| = 1$, and thus Theorem 3.1 becomes [4, Theorem 2.4]. Further, note that when T is J-unitary then (3.7) and (3.9) imply $\widetilde{H} = H$; so it comes as no surprise that the right-hand side of (3.18) is zero, as it should. The appearances of norms of those V matrices (with or without hats or tildes) in the bound (3.18), as well as in those below, is a nuisance; whether they could be removed is not clear to us, but we suspect that they cannot.

In general when A is indefinite, it is not clear at all how big $\|\widehat{V}_2\|$ and $\|\widetilde{V}_1\|$ may get. Also we would like to make $\|\widehat{V}_2\|$ disappear from the bound since it corresponds to the intermediate $\widehat{H}$. For practical purpose, $\|J\widetilde{T}^* - \widetilde{T}^{-1}J\|_{\mathrm{F}}$ is not immediately available and will likely be bounded in terms of norms of E (and thus of $\|\Delta A\|_{\mathrm{F}}$). We shall now deal with these issues. For the ease of presentation, define

$$\delta = \|A^{-1}\|\|\Delta A\|, \quad \delta_{\mathrm{F}} = \|A^{-1}\|\|\Delta A\|_{\mathrm{F}}. \tag{3.19}$$

It can be seen that $\|E\| \leqslant \delta$ and $\|E\|_{\mathrm{F}} \leqslant \delta_{\mathrm{F}}$.

3.1. Bounding $\|J\widetilde{T}^* - \widetilde{T}^{-1}J\|_{\mathrm{F}}$

We shall present a couple of lemmas in which the factor $\|J\widetilde{T}^* - \widetilde{T}^{-1}J\|_{\mathrm{F}}$ in the right-hand side of (3.18) will be bounded in terms of $T - T^{-1}$, ΔA, and A. Using (3.16), $T^* = JTJ$, $W^{-1} = W$, and $W^* = JWJ$, we have

$$\begin{aligned} J\widetilde{T}^* - \widetilde{T}^{-1}J &= TWT^{-1}J - T^{-1}WTJ \\ &= (TW(T^{-1} - T) + (T - T^{-1})WT)J. \end{aligned}$$

Thus we have the following bound

$$\left\|J\widetilde{T}^* - \widetilde{T}^{-1}J\right\|_{\mathrm{F}} \leqslant 2\|W\|\|T\|\|T^{-1} - T\|_{\mathrm{F}}. \tag{3.20}$$

Write $T = I + \varGamma$. We have by (3.8)

$$\begin{aligned} \|\varGamma\| &\leqslant \sum_{n=1}^{\infty} \frac{(2n-1)!!}{2^n n!}\|E\|^n \leqslant \frac{1}{2}\|E\| \sum_{n=1}^{\infty} \|E\|^{n-1} \\ &= \frac{1}{2}\frac{\|E\|}{1 - \|E\|} \leqslant \frac{1}{2}\frac{\delta}{1-\delta}, \end{aligned} \tag{3.21}$$

$$\|\varGamma\|_{\mathrm{F}} \leqslant \frac{1}{2}\frac{\|E\|_{\mathrm{F}}}{1 - \|E\|} \leqslant \frac{1}{2}\frac{\delta_{\mathrm{F}}}{1-\delta}. \tag{3.22}$$

So if $\delta < 2/3$, $\|\Gamma\| < 1$ which implies $T^{-1} = I - \Gamma + \Gamma^2 - \Gamma^3 + \cdots$, and thus

$$\begin{aligned}\|T^{-1} - T\|_{\mathrm{F}} &\leqslant \|\Gamma\|_{\mathrm{F}} + \frac{\|\Gamma\|_{\mathrm{F}}}{1 - \|\Gamma\|} \\ &= \|\Gamma\|_{\mathrm{F}} \frac{2 - \|\Gamma\|}{1 - \|\Gamma\|} \\ &\leqslant \frac{1}{2} \frac{4 - 5\delta}{(1-\delta)(2-3\delta)} \delta_{\mathrm{F}} \\ &\leqslant \frac{2}{2 - 3\delta} \delta_{\mathrm{F}}. \end{aligned} \tag{3.23}$$

An immediate consequence of Theorem 3.1, (3.15), (3.20), and (3.23) is the following corollary.

Corollary 3.1. *To the conditions of Theorem* 3.1 *add this*: $\delta \equiv \|A^{-1}\|\|\Delta A\| < 2/3$. *Then*

$$\frac{1}{2}\left\| \sin 2\Theta(U_1, \widetilde{U}_1)\right\|_{\mathrm{F}} \leqslant \|V\|^2 \|\widetilde{V}_1\| \|\widehat{V}_2\| \frac{\|T\|_2 \|T - T^{-1}\|_{\mathrm{F}}}{\widetilde{\eta}_\chi}, \tag{3.24}$$

$$\leqslant \|V\|^2 \|\widetilde{V}_1\| \|\widehat{V}_2\| \frac{\varepsilon}{\widetilde{\eta}_\chi}, \tag{3.25}$$

where

$$\varepsilon = \frac{2 - \delta}{(1-\delta)(2-3\delta)} \delta_{\mathrm{F}} = \delta_{\mathrm{F}} + \mathcal{O}(\delta_{\mathrm{F}}^2). \tag{3.26}$$

Note that bounds (3.24) and (3.25) are proper generalizations of [4, (2.35)], since in the positive definite case V, $\widetilde{V}$ and $\widehat{V}$ are unitary and thus disappear from these inequalities altogether.

3.2. Bounding $\|\widetilde{V}\|$ and $\|\widehat{V}\|$ in terms of $\|V\|$

We shall now bound $\|\widetilde{V}\|$ and $\|\widehat{V}\|$ in terms of $\|V\|$. For this we will need the following lemma.

Lemma 3.1. *Let V and $\widetilde{V}$ be the canonical eigenvector matrices of the pairs $\{G^*G, J\}$ and $\{T^*G^*GT, J\}$, respectively. Write $\Gamma = T - I$ and define $\gamma = \|\Gamma\|_{\mathrm{F}}/(1 - \|\Gamma\|)$.*

$$If\ \|\Gamma\| < 1\ and\ \gamma\|V\|^2 < \frac{1}{4},\ then\ \|\widetilde{V}\| \leqslant \frac{\|V\|}{\sqrt{1 - 4\gamma\|V\|^2}}. \tag{3.27}$$

Proof. Eq. (3.27) follows from [8, Lemma 5] (see [8, Lemma 4] or [6, Lemma 1]). □

It is reasonable to expect that $\|\Gamma\|_{\mathrm{F}}$ in defining γ in this lemma should be replaced by $\|\Gamma\|$, the spectral norm of Γ. But we are unable to prove this.

To use Lemma 3.1, we shall interpret that V, $\widetilde{V}$, and $\widehat{V}$ assigned above are the canonical eigenvector matrices of the definite pairs $\{G^*G, J\}$, $\{\widetilde{G}^*\widetilde{G}, J\}$, and $\{\widehat{G}^*\widehat{G}, J\}$, respectively (see (3.5), (3.11), and (3.16) for the assignments of G, $\widetilde{G}$, and $\widehat{G}$). By (3.22),

$$\frac{\|\Gamma\|_{\mathrm{F}}}{1-\|\Gamma\|} \leqslant \frac{1}{2-3\delta}\delta_{\mathrm{F}} \equiv \alpha. \tag{3.28}$$

Lemma 3.1 applied to $\{G^*G, J\}$ and $\{\widetilde{G}^*\widetilde{G}, J\}$ yields

$$\|\widetilde{V}\| \leqslant \frac{\|V\|}{\sqrt{1-4\alpha\|V\|^2}} \quad \text{if } 4\alpha\|V\|^2 < 1 \text{ and } \delta < 2/3. \tag{3.29}$$

Note by (3.11) and (3.16) and $\widehat{G} = GWT$, thus $\widehat{G}^*\widehat{G} = T^*W^*G^*GWT$ and

$$GW = U|\Lambda|^{1/2}(W^{-1}V)^{-1}, \tag{3.30}$$

by (3.15). Since both W^{-1} and V are J-unitary, $W^{-1}V$ is also J-unitary. Eq. (3.30) is the HSVD of GW and consequently $W^{-1}V$ is the canonical eigenvector matrix of the pair $\{W^*G^*GW, J\}$. It follows from (3.14) that

$$W^{-1}V = V\begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{bmatrix},$$

and thus $\|W^{-1}V\| = \|V\|$. Lemma 3.1 applied to the pairs $\{W^*G^*GW, J\}$ and $\{\widehat{G}^*\widehat{G}, J\}$ (recall $\widehat{G} = GWT$) implies

$$\|\widehat{V}\| \leqslant \frac{\|V\|}{\sqrt{1-4\alpha\|V\|^2}} \quad \text{if } 4\alpha\|V\|^2 < 1 \text{ and } \delta < 2/3. \tag{3.31}$$

Corollary 3.2. *To the conditions of Theorem* 3.1 *add these*:

$$\delta < 2/3, \quad \alpha \equiv \frac{1}{2-3\delta}\delta_{\mathrm{F}} \leqslant \frac{1}{4\|V\|^2}.$$

Then

$$\frac{1}{2}\|\sin 2\Theta(U_1, \widetilde{U}_1)\|_{\mathrm{F}} \leqslant \frac{\|V\|^4}{1-4\alpha\|V\|^2}\frac{\|T\|_2\|T-T^{-1}\|_{\mathrm{F}}}{\widetilde{\eta}_\chi}, \tag{3.32}$$

$$\leqslant \frac{\|V\|^4}{1-4\alpha\|V\|^2}\frac{\varepsilon}{\widetilde{\eta}_\chi}, \tag{3.33}$$

where ε is defined by (3.26).

Proof. By (3.28), γ in Lemma 3.1 satisfies $\gamma \leqslant \alpha$. The inequality (3.32) follows by inserting (3.27) and (3.31) into (3.24) and (3.33) follows by inserting (3.27) and (3.31) into (3.25). $\square$

3.3. Bounding V

The bounds (3.32) and (3.33) contain an additional factor which depends on J-unitary matrix V whose norm may be big. Here we shall show that $\|V\|$ is of modest magnitude for special but interesting matrices. In [8] two classes of so-called *well-behaved matrices* for which $\kappa(V) \equiv \|V\|\|V^{-1}\|$ can be bounded in a satisfactory way are defined. The first class consists of *scaled diagonal dominant* (SDD) *matrices*, and the second class consists of *quasi-definite matrices*. We shall now review bounds in [8], as well as derive new and improved ones, for both classes of matrices.

A matrix H is SDD, if it can be written as $H = D(J + N)D$ with diagonal positive definite D, $J = \mathrm{diag}(\pm 1)$, and $\|N\| < 1$. The following theorem was proved in [8].

Theorem 3.2. *Let $H = D(J + N)D$ be SDD. Then $\|V\| \leqslant \sqrt{n(1 + \|N\|)/(1 - \|N\|)}$.*

This bound depends on the square root of the dimension. Note that

$$H = D(I + NJ)^{1/2} J[(I + NJ)^{1/2}]^* D = \widetilde{B} J \widetilde{B}^*,$$

where $\widetilde{B} = D(I + NJ)^{1/2}$. Thus, we can interpret $\widetilde{B}$ as obtained by multiplicatively perturbing D, and this way V (the canonical eigenvector matrix of the pair $\{[(I + NJ)^{1/2}]^* D^2 (I + NJ)^{1/2}, J\}$) is resulted from perturbing I (canonical eigenvector matrix of the pair $\{D^2, J\}$). Lemma 3.1 applied to the two pairs yields a theorem as follows.

Theorem 3.3. *Let $H = D(J + N)D$ be SDD. If $\|N\|_{\mathrm{F}} < 2/7$, we have*

$$\|V\|^2 \leqslant \frac{1}{1 - 4\gamma}, \tag{3.34}$$

where $\gamma = \|N\|_{\mathrm{F}}/(2 - 3\|N\|)$.

Proof. Write $I + \Gamma = (I + NJ)^{1/2}$. Analogously to (3.22), we have $\|\Gamma\|_{\mathrm{F}} \leqslant \frac{1}{2}\|N\|_{\mathrm{F}}/(1 - \|N\|)$. Thus if $\|N\|_{\mathrm{F}} < 2/7$, $\|\Gamma\|_{\mathrm{F}}/(1 - \|\Gamma\|) \leqslant \gamma < 1/4$, and then (3.27) implies (3.34). □

The bound in Theorem 3.3 does not explicitly depend on n while the bound in Theorem 3.2 does.

Next we shall extend the bound (3.34) to a larger class of matrices, containing SDD matrices. We say a Hermitian matrix H *is block scaled diagonally dominant* (BSDD) if it admits

$$H = D_{\mathrm{b}}^* (J_{\mathrm{b}} + N_{\mathrm{b}}) D_{\mathrm{b}}, \tag{3.35}$$

where $D_{\mathrm{b}} = D_1 \oplus D_2 \oplus \cdots \oplus D_k$ and D_i is nonsingular, $J_{\mathrm{b}} = J_1 \oplus J_2 \oplus \cdots \oplus J_k$, with $J_i = I$ or $J_i = -I$, and $\|N_{\mathrm{b}}\|_{\mathrm{F}} \leqslant 1$. A BSDD H as just described can be rewritten as

$$H \equiv BJ_bB^*, \quad \text{where } B = D_b^*(I + N_bJ_b)^{1/2}. \tag{3.36}$$

Using similar approach as above for SDD matrices we have the following bound.

Theorem 3.4. *Let $H = D_b^*(J_b + N_b)D_b$ be BSDD. If $\|N_b\|_F < 2/7$, we have*

$$\|V\|^2 \leqslant \frac{1}{1-4\gamma_b}, \tag{3.37}$$

where $\gamma_b = \|N_b\|_F/(2-3\|N_b\|)$.

Lastly we consider so-called quasi-definite matrices. A Hermitian matrix H is said to be a *quasi-definite* if there exists a permutation matrix P such that

$$H_q \equiv P^THP = \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^* & -H_{22} \end{bmatrix},$$

where H_{11} and H_{22} are positive definite. The following theorem was proved in [8].

Theorem 3.5. *Let H be quasi-definite as just described. Then*

$$\kappa(V) \leqslant n \max\{\|A_{11}\| + \|A_{12}A_{22}^{-1}A_{12}^*\|,\ \|A_{22}\| + \|A_{12}^*A_{11}^{-1}A_{12}\|\}, \tag{3.38}$$

where

$$H_q \equiv DAD = D\begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & -A_{22} \end{bmatrix}D,$$

$D = \operatorname{diag}(H_{11}^{1/2}, H_{22}^{1/2})$, *and V is the canonical eigenvector matrix of the pair $\{G^*G, J\}$, $J_{ii} = \operatorname{sign}(H_{ii})$ and $G = DF$ such that $H_q = GJG = DFJF^*D$ (that is, $A = FJF^*$).*

Theorem 3.4 also applies to the current case if H_{12} is "small" enough. To do so, we let $H_{11} = L_1L_1^*$ and $H_{22} = L_2L_2^*$ be, e.g., Cholesky factorizations, and write

$$H_q = L_q(J + N_q)L_q^*,$$

where

$$L_q = \begin{bmatrix} L_1 & \\ & L_2 \end{bmatrix},\quad N_q = \begin{bmatrix} 0 & L_1^{-1}H_{12}L_2^{-*} \\ L_2^{-1}H_{12}^*L_2^{-*} & 0 \end{bmatrix},\quad J = \begin{bmatrix} I & \\ & -I \end{bmatrix}.$$

Note that if $\|N_q\|$ is small enough we can apply (3.27). Write $\gamma_q = \|N_q\|_F/(2 - 3\|N_q\|)$. If $\|L_1^{-1}H_{12}L_2^{-*}\|_F < 2/(4\sqrt{2}+3)$ which implies $\gamma_q < 1/4$, from (3.37) it follows

$$\|V\|^2 \leqslant \frac{1}{1-4\gamma_q}. \tag{3.39}$$

We have so far considered three related classes, and all fall into the category of BSDD matrices. So we shall present a theorem as a corollary to Theorem 3.1 for BSDD matrices.

Theorem 3.6. *Let $H = D_{\mathrm{b}}^* A D_{\mathrm{b}}$ be BSDD matrix where $A = J_{\mathrm{b}} + N_{\mathrm{b}}$, $D_{\mathrm{b}} = D_1 \oplus \cdots \oplus D_k$ and $J_{\mathrm{b}} = J_1 \oplus \cdots \oplus J_k$ are nonsingular with $J_i = I$ or $J_i = -I$, for $i = 1, \ldots, k$. Suppose H is perturbed to $\widetilde{H} = D^*(A + \Delta A)D$. Set*

$$\alpha_{\mathrm{b}} \equiv \frac{\delta_{\mathrm{F}}}{2 - 3\delta} = \frac{\left\|(J_{\mathrm{b}} + N_{\mathrm{b}})^{-1}\right\| \|\Delta A\|_{\mathrm{F}}}{2 - 3\left\|(J_{\mathrm{b}} + N_{\mathrm{b}})^{-1}\right\| \|\Delta A\|} \quad \textit{and} \quad \gamma_{\mathrm{b}} \equiv \frac{\|N_{\mathrm{b}}\|_{\mathrm{F}}}{2 - 3\|N_{\mathrm{b}}\|}.$$

If $\|N_{\mathrm{b}}\|_{\mathrm{F}} < 2/7$ and if $4\alpha_{\mathrm{b}}/(1 - 4\gamma_{\mathrm{b}}) < 1$, then

$$\frac{1}{2}\left\| \sin 2\Theta(U_1, \widetilde{U}_1)\right\|_{\mathrm{F}} \leqslant \frac{1}{(1 - 4\gamma_{\mathrm{b}})(1 - 4(\gamma_{\mathrm{b}} + \alpha_{\mathrm{b}}))} \frac{\|T\|_2 \|T - T^{-1}\|_{\mathrm{F}}}{\widetilde{\eta}_\chi}, \tag{3.40}$$

$$\leqslant \frac{1}{(1 - 4\gamma_{\mathrm{b}})(1 - 4(\gamma_{\mathrm{b}} + \alpha_{\mathrm{b}}))} \frac{\varepsilon}{\widetilde{\eta}_\chi}, \tag{3.41}$$

where ε is defined by (3.26).

Proof. The upper bounds are obtained by bounding α and $\|V\|$ as in Corollary 3.2. For BSDD matrices we have $A^{-1} = (J + N_{\mathrm{b}})^{-1}$. Further since $\|\alpha_{\mathrm{b}}\|_{\mathrm{F}} < 2/7$, from (3.37) it follows that $\|V\|^2 \leqslant 1/(1 - 4\gamma_{\mathrm{b}})$, and thus $4\alpha_{\mathrm{b}}\|V\|^2 < 1$ which allows us to use the bounds (3.32) and (3.33). It can be seen that

$$\frac{\|V\|^4}{1 - 4\alpha_{\mathrm{b}}\|V\|^2} \leqslant \frac{1}{(1 - 4\gamma_{\mathrm{b}})(1 - 4(\gamma_{\mathrm{b}} + \alpha_{\mathrm{b}}))},$$

which together with (3.32) and (3.33) yield (3.40) and (3.41), respectively. □

4. A numerical example

All computations in the following example were performed within MATLAB environment. Double precision arithmetic was used throughout. In the example that follows, input data are exactly the same as shown in decimal; but only five decimal digits for all outputs are printed here to save space. Let $H = D^*AD$, where

$$D = \begin{bmatrix} L_1 & \\ & L_2 \end{bmatrix}, \quad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21}^* & A_{22} \end{bmatrix},$$

$$L_1 = \begin{bmatrix} 6.7676\mathrm{e}+2 & 4.4692\mathrm{e}+2 & 7.6564\mathrm{e}+2 \\ 0 & 2.3767\mathrm{e}-1 & 8.5502\mathrm{e}-2 \\ 0 & 0 & 1.2205\mathrm{e}-1 \end{bmatrix},$$

$$L_2 = \begin{bmatrix} 1.4078\mathrm{e}+3 & -4.4582\mathrm{e}+2 & -2.1765\mathrm{e}+3 \\ 0 & 3.4030\mathrm{e}+3 & -5.5373\mathrm{e}+3 \\ 0 & 0 & 4.8642\mathrm{e}+1 \end{bmatrix},$$

$$A_{11} = I,\ A_{22} = -I,$$

$$A_{12} = \begin{bmatrix} 4.6675\text{e}-3 & 3.9688\text{e}-3 & 1.0076\text{e}-3 \\ 9.7469\text{e}-2 & 4.7937\text{e}-2 & 3.4591\text{e}-2 \\ 2.7080\text{e}-5 & -2.9377\text{e}-5 & -5.8637\text{e}-4 \end{bmatrix}.$$

Let

$$\Delta A = \begin{bmatrix} 2.2\text{e}-4 & 1.1\text{e}-5 & 2.4\text{e}-5 & 3.8\text{e}-6 & 3.2\text{e}-6 & 1.3\text{e}-5 \\ 1.1\text{e}-5 & 8.4\text{e}-6 & 3.4\text{e}-4 & 3.0\text{e}-6 & 3.5\text{e}-5 & 9.9\text{e}-6 \\ 2.4\text{e}-5 & 3.4\text{e}-4 & 6.7\text{e}-5 & 1.3\text{e}-8 & -2.1\text{e}-6 & -2.2\text{e}-7 \\ 3.8\text{e}-6 & 3.0\text{e}-6 & 1.3\text{e}-8 & -9.2\text{e}-5 & 3.4\text{e}-4 & 4.7\text{e}-5 \\ 3.2\text{e}-6 & 3.5\text{e}-5 & -2.1\text{e}-6 & 3.4\text{e}-4 & -4.7\text{e}-5 & 1.9\text{e}-5 \\ 1.3\text{e}-5 & 9.9\text{e}-6 & -2.2\text{e}-7 & 4.7\text{e}-5 & 1.9\text{e}-5 & -2.5\text{e}-4 \end{bmatrix}.$$

Spectrum of matrix $\widetilde{H} = D^*(A + \Delta A)D$ is

$$\begin{aligned}\lambda(\widetilde{H}) = \{&-4.5142\text{e}+7, -4.0159\text{e}+6, -2.9967\text{e}+2, 7.6909\text{e}-3, \\ &4.0809\text{e}-2, 1.2443\text{e}+6\}.\end{aligned}$$

If we write $A = Q|\Omega|^{1/2}J|\Omega|^{1/2}Q^*$ then $H = GJG^*$ (similarly for perturbed quantities). Now applying the HSVD [1] to $\widetilde{G}$ we obtain $\widetilde{G} = \widetilde{U}|\widetilde{\Lambda}|^{1/2}\widetilde{V}^{-1}$, where

$$\widetilde{U} = \begin{bmatrix} 6.0678\text{e}-1 & -4.4527\text{e}-1 & -6.5844\text{e}-1 & -4.9637\text{e}-4 & 8.8632\text{e}-4 & 1.8055\text{e}-4 \\ 4.0071\text{e}-1 & 8.8676\text{e}-1 & -2.3040\text{e}-1 & -3.3034\text{e}-4 & 5.9491\text{e}-4 & -3.7053\text{e}-4 \\ 6.8647\text{e}-1 & -1.2404\text{e}-1 & 7.1650\text{e}-1 & -5.6247\text{e}-4 & 1.0062\text{e}-3 & 3.2199\text{e}-5 \\ 1.0515\text{e}-3 & 3.0853\text{e}-4 & 7.8402\text{e}-6 & 5.5603\text{e}-2 & -6.7732\text{e}-1 & 7.3358\text{e}-1 \\ -5.8103\text{e}-4 & 2.3473\text{e}-4 & 5.9004\text{e}-6 & 4.7068\text{e}-1 & 6.6574\text{e}-1 & 5.7901\text{e}-1 \\ -1.1751\text{e}-3 & 1.4252\text{e}-4 & 3.5703\text{e}-6 & -8.8055e-1 & 3.1309\text{e}-1 & 3.5582\text{e}-1 \end{bmatrix},$$

$$\widetilde{V} = \begin{bmatrix} 9.9786\text{e}-1 & -6.0033\text{e}-2 & 2.6174\text{e}-2 & -3.2594\text{e}-4 & -1.4212\text{e}-3 & 1.5757\text{e}-3 \\ 5.6316\text{e}-2 & 9.9210\text{e}-1 & 1.2569\text{e}-1 & -3.5942\text{e}-2 & 4.0352\text{e}-2 & -1.7707\text{e}-2 \\ -3.3464\text{e}-2 & -1.2391\text{e}-1 & 9.9173\text{e}-1 & 1.8614\text{e}-3 & -2.0045\text{e}-3 & 1.1523\text{e}-3 \\ -1.9456\text{e}-3 & -5.6730\text{e}-2 & -4.1672\text{e}-3 & 6.0271\text{e}-1 & -7.3069\text{e}-1 & 3.2568\text{e}-1 \\ -4.4352\text{e}-4 & -1.3248\text{e}-3 & 6.3785\text{e}-5 & 5.5096\text{e}-1 & 6.7381\text{e}-1 & 4.9236\text{e}-1 \\ 1.7198\text{e}-3 & 1.6388\text{e}-3 & 3.5202\text{e}-4 & -5.7834\text{e}-1 & -1.1706\text{e}-1 & 8.0736\text{e}-1 \end{bmatrix}.$$

We shall derive a bound for $\|\sin 2\Theta(U_1, \widetilde{U}_1)\|_{\text{F}}$, where U_1 and $\widetilde{U}_1$ contain eigenvectors corresponding to eigenvalues $\lambda_1 = 1.2440\text{e}+6$, $\lambda_2 = 7.6900\text{e}-3$, and $\widetilde{\lambda}_1 = 1.2443\text{e}+6$, $\widetilde{\lambda}_2 = 7.6909\text{e}-3$, respectively. We have

$$\alpha_{\text{b}} = 8.7366\text{e}-004, \quad \gamma_{\text{b}} = 9.1307\text{e}-002, \quad \widetilde{\eta}_\chi = 1.8694.$$

Eq. (3.41) of Theorem 3.6 gives

$$\frac{1}{2}\|\sin 2\Theta(U_1, \widetilde{U}_1)\|_{\text{F}} \leqslant 1.0333\text{e}-003,$$

in comparison to $\|\sin 2\Theta(U_1, \widetilde{U}_1)\|_{\text{F}} = 3.5304\text{e}-004$.

[1] A MATLAB function to compute HSVDs is available from the first author upon request.

Note that the absolute gap is $3.3118\mathrm{e}-2$ and $\|\Delta H\| = 7.2828\mathrm{e}+3$, and thus the classical Davis–Kahan $\sin 2\Theta$ theorem produces

$$\frac{1}{2}\left\| \sin 2\Theta(U_1, \widetilde{U}_1)\right\|_{\mathrm{F}} \leqslant 2.2\mathrm{e}+5$$

which is too big to be useful. Also note, since $\|N_{\mathrm{b}}\| = 0.11415$ the bound (3.39) yields that $\|V\| \leqslant 1.2551$, a quite accurate estimation of $\|V\| = 1.059$.

Acknowledgements

The authors would like to thank Professor B.N. Parlett's and anonymous referees' constructive criticism that greatly improves our presentation.

References

[1] C. Davis, W.M. Kahan, The rotation of eigenvectors by a perturbation III, SIAM J. Numer. Anal. 7 (1970) 1–46.
[2] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1990.
[3] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.
[4] R.-C. Li, Relative perturbation theory: (iv) sin 2Θ theorems, Linear Algebra Appl. 311 (2000) 45–60.
[5] R. Onn, A.O. Steinhardt, A. Bojanczyk, Hyperbolic singular value decompositions and applications, IEEE Trans. Acoustics, Speech, Signal Process. (1991) 1575–1588.
[6] I. Slapničar, N. Truhar, Relative perturbation theory for hyperbolic eigenvalue problem, Linear Algebra Appl. 309 (2000) 57–72.
[7] G.W. Stewart, J.-G. Sun, Matrix Perturbation Theory, Academic Press, Boston, 1990.
[8] N. Truhar, I. Slapničar, Relative perturbation bound for invariant subspaces of Hermitian matrices, Linear Algebra Appl. 301 (1999) 171–185.
[9] H. Zha, A note on the existence of the hyperbolic singular value decomposition, Linear Algebra Appl. 240 (1996) 199–205.